记忆的前测效应对老年人学习新事物的作用[1]

王堂生[a, c]; 杨春亮[b, 2];钟年[c]

[a] 武汉理工大学，武汉，430070

[b] 北京师范大学，北京，100875

[c] 武汉大学，武汉，430072

摘要 为检验前置测试是否可以有效地提升老年人学习新事物的能力，参与实验的老年人被随机分为有前置测试的实验组和无前置测试的对照组，在实验 1 中学习了 5 列单词，在实验 2 中记忆了 5 列日用品，在实验 3 中学习了 4 段讲座视频。结果表明，实验组老年人对新事物的记忆成绩均高于对照组的老年人。结论认为，前置测试能够显著提升老年人学习新事物的能力，干扰降低与学习参与相结合的“加减”理论能更好地解释提升老年人记忆的前测效应。

关键词 前置测试；前测效应；老年人；学习参与；干扰降低

1 引言

当前许多国家正迈入加速老龄化的社会，60 岁以上的人比重越来越大，他们的工作和生活状况在这些国家都变得非常重要，而对其认知能力的研究、评估和改善，则事关该年龄群体的个人生活、社会制度和国家政策等多方面的重要决定。

在老年人认知的研究中，记忆老化可能是至关重要的内容，其中大部分研究发现是负面的，如工作记忆资源有限（Whitebourne & Whitebourne, 2014），老年痴呆风险增高（Hafkemeijer, Van & Rombouts,2012），以及情节记忆(Old & Naveh-Benjamin, 2008)、往事记忆(Piolino, Desgranges, Benali, & Eustache, 2002)、未来记忆(Farrimond, Knight, & Titov, 2003)等方面均存在随着老化而能力下降的现象。

根据既往的研究，尽管老化对记忆可能有诸多消极的影响，但也有一些积极的发现，老年人在语义（Wiggs, Weisberg, & Martin, 2006）、自我经历（Gluck & Bluck, 2007）等方面的记忆力极少下降，而且并非每一个人在老化进程中都以同样的方式受到影响

[1] 中央高校基本科研业务费专项资金资助（项目批准号：2020VI022）

[2] 通讯作者，chunliang.yang@bnu.edu.cn

（Whitebourne & Whitebourne， 2014），其中令人欣慰的发现是，记忆训练有助于老年人认知功能的改善（Craik & Rose, 2012）。

本研究主要探索前测效应是否有助于改善老年人的记忆成绩, 从而为进一步改善老年人的学习和记忆能力奠定基础。

大量的研究表明，与学习后重学或未做任何事情相比，对已经学过的信息进行测试有助于学习和记住新信息，这被称为前测效应（forward testing effect， Pastötter & Bäuml, 2014; Yang, Potts, & Shanks, 2018）。正如 Yang 等人（2018）所总结的那样，对于不同类型的事物例如词汇(Szpunar, McDermott, & Roediger, 2008)、图片(Pastötter, Weber, & Bäuml, 2013)、外语翻译词汇配对（Yang， Potts， & Shanks， 2017）、面孔-姓名配对（Weinstein, McDermott, & Szpunar, 2011）、文本段落（Wissman, Rawson, & Pyc, 2011）、演讲视频（Szpunar, Khan, & Schacter, 2013）、图画（Yang & Shanks， 2018），以及其他学习材料而言，前测效应是一种稳定而可靠的现象。前测效应并不仅存在于有教师指导下的学习状态，并且可推及到自行规划的学习状态（Yang, Chew, Sun, & Shanks, 2019）。

Szpunar 等人(2008) 进行了一项经典的研究，发现了对简单事物进行学习时的前测效应。他们指导两组参与者（大学生）学习 5 列单词，每组 18 个，参与者逐词逐列学习。实验组参与者在学习完每一列单词之后都会分别进行临时测试（interim test recall）。控制组的参与者则不参加第 1 至第 4 列单词的临时测试，只参加第 5 个单词表的临时测试。实验结果表明，在第 5 列单词的临时测试中，实验组能够准确记住的单词量是控制组的两倍，这说明了学习简单事物时的前测效应。此外，在第 5 列单词的临时测试中，由于前摄干扰效应(proactive inference,PI)，控制组出现的错误是实验组的十几倍(要求参与者回忆第 5 列的单词时，他们错误地回忆起了第 1 至第 4 列的单词)。关于学习简单事物的前测效应最近已有多篇文献发表(如：Nunes & Weinstein,2012；Pastötter, Schicker, Niedernhuber, & Bäuml, 2011；Weinstein, Gilmore, Szpunar, & McDermott, 2014; Yang， et al.， 2017）。Szpunar 等人（2008) 提出干扰降低理论（Release from PI），旨在解释简单事物学习的前测效应。根据该理论的解释，分列回忆策略实际上改变了目标词汇的回忆背景，有利于对不同词列进行分隔处理，降低了前摄效应的干扰。实验组因为进行了临时测试，该情节记忆使得新事物的记忆具有较少的背景事物，这可以降低前摄干扰效应，从而促进了对目标事物的回忆（如第 5 列中的单词）。

在上述简单事物的学习实验中，相对于第 5 列的临时测试，前 4 列的临时测试被称为前置测试（Interpreted testing），它们不仅可以促进对新的简单事物的学习，还有助于对演讲

视频（Jing, Szpunar, & Schacter, 2016）、文章段落(Wissman, et al., 2011）等类似的复杂事物的学习。Szpunar 等人 (2013)曾指导 3 组参与者对 4 段演讲视频进行逐段学习。实验组参与者在学完每一段视频后都进行了临时测试, 另外两组则只在学完最后一段视频后进行临时测试——前 3 段视频学完之后要么重学一遍（重学组），要么演算数学题（演算组）。3 组参与者都可以做笔记，同时要求他们报告在观看视频时他们的思维是否处于工作状态，结果发现实验组对于测试后学习的视频内容记忆成绩非常好, 而演算组和重学组的记忆成绩显著低于实验组，这证明了对复杂事物的学习的前测效应。Szpunar 等人（2013）还观察到实验组比其它小组做了更多的笔记，而报告的思绪游移（Mind Wandering，MW）也更少。对于复杂事物学习的前测效应被近期许多研究反复证明（Schacter & Szpunar, 2015; Szpunar, Jing, & Schacter, 2014）。

在简单事物的学习中，干扰降低理论在很大程度上可以解释前测效应。然而，这种理论难于解释对于复杂事物学习的前测效应，因为对于复杂事物而言，前摄干扰效应（PI）有限，前测效应似乎不会发生（Wissman, et al., 2011）。Szpunar 等人（2013）的干扰降低理论认为, 复杂事物学习的前测效应是基于这样的事实——前置测试降低了与任务无关的思绪游移（MW）。不过，Yang 等人(2017)观察到，在无临时测试时的学习过程中，参与者的学习动力（学习时间）不断下降， 如果插入了阶段性的临时测试，他们的学习动力可以一直保持。Pastötter 等人于 2011 年发现，在学习过程中，学习者对记忆任务的关注有所降低，但是进行临时测试时这种降低又得以恢复。所有这些发现都支持学习参与 (learning engagement, LE）理论，即对复杂事物学习的前测效应是指在周而复始的学习过程中，前置测试使学习参与的降低得到了回升，这有助于新的复杂事物的学习。

总而言之，人们对于不同类事物的学习都表现出稳定而可靠的前测效应。简单事物学习的前测效应可能是因为前置测试抑制了记忆的前摄干扰（PI），复杂事物的学习的前测效应则可能是因为前置测试增强了学习参与（LE）的动机。有一点很重要，就是要注意两种机制（降低干扰和增强参与）并不相互排斥，两种机制都可能对简单事物和复杂事物学习的前测效应具有促进作用。此外，前置测试对学习过程中的思绪游移（MW）现象也可能具有抑制作用。

前测效应是否可以被推论到老年人？尽管老年人的记忆增强机制可能与年轻人存在差异, 但是有更多的理由预测前置测试可以作为一种有效的弥补策略有助于老年人学习和记忆新的信息。本研究假设：在即将开展的三个实验中，无论是对简单事物的学习还是对复杂事物的学习，前置测试均会显著提高老年人对新的材料的记忆成绩，即：前置测试会促进老年

人对新的学习材料的记忆成绩（假设一），但其驱动机制并不局限于前摄干扰效应的抑制作用（假设二）。

已经有若干相关的研究支持肯定的假设，前测效应在其中获得了稳定而可靠的证据。Pastötter 等人(2013)已经证实前测效应可以推论到大脑创伤病人, Pastötter 和 Bäuml (2019) 也将前测效应推论到中老年人对新的简单事物的学习中, 主要探索了前置测试通过抑制中老年人前摄干扰效应而促进了新的简单事物（4 至 8 个字母的德语名词）的学习，不过对于涉及复杂事物学习的动机参与（LE）、思绪游移（MW）等心理作用机制则没有进行探索。而与本文研究假设相反的推测是, 前测效应无法将这种结果推论到 60 岁以上的老年人群体中, 既包括对简单事物又包括对复杂事物的学习。在简单事物学习方面，Aslan 和 Bäuml（2015）就儿童对简单事物的学习的前测效应的研究似乎为这种相反的推测提供了证据。他们发现，对于年轻人和高龄儿童来说，临时测试基本上可以避免 PI 的生成，但是对于低龄儿童来说，临时测试不能降低其 PI。Aslan 和 Bäuml（2015）认为，低龄儿童学习简单事物时之所以没有前测效应可能是因为低龄儿童难以控制 PI。与低龄儿童的情况相似，老年人在抑制 PI 方面遭遇的困难大于年轻成人（Ikier & Hasher,2006; Ikier, Yang, & Hasher, 2008）。因此，有理由推断高龄儿童和年轻成人对于学习新事物的前测效应并不能直接推论到老年人。如果老年人在学习复杂事物时存在前测效应，由于老年人在执行不同类型任务时出现的思绪游移（MW）较少，如阅读文章段落时（Krawietz, Tamplin, & Radcansky, 2012），就将使干扰降低理论的解释面临困难，而老人们对相关任务能够持续保持注意（SART;Jackson & Balota, 2012），这又使得学习参与理论的解释面临挑战。当然，这些相反的假设证据相对较少，并且都是较为间接的推测。

2 研究方法、过程与结果

2.1 实验 1

2.1.1 参与者

本次实验是一次线上学习，参与者是通过富饶学术[3]招募的，实验者对参与者设置了如下的条件限制：母语为英语，年龄超过 60 岁，国籍为英国，当前居住地为英国。在富饶学术的 44327 个人中有 127 个人符合条件。

样本的大小是根据 Yang 等人(2017)实验 4 中的效应量（Cohen’s $d = 1.12$）决定的，样本量要求在每组的 14 个参与者中观察到显著的前测效应（$\alpha = 0.05$; power = 0.80）。为了平

[3] 注：Prolific Academic 参见网址 https://www.prolific.ac/

衡实验材料，最终每组参与者的样本数确定为 15 个。共招募了 30 名老年人（17 名为女性，13 名男性），被随机分入两组，一组为测试组，一组为演算组（对照组），每组 15 人。有三个参与者报告曾经罹患过精神类的疾病（临床诊断），测试组有 1 名抑郁症患者，演算组有 1 名抑郁症患者和 1 名抑郁焦虑混合症患者。排除这三名参与者的数据不影响整体结果的模式，所以本文保留了他们的数据。此外还对参与者其他的人口统计学数据（年龄、教育程度和健康状况）进行了调查。在实验最开始，实验者明确地告诉参与者不要做笔记以辅助他们的记忆。在实验最后，所有的参与者被要求报告是否做了笔记以辅助他们的记忆。所有的参与者都报告没有记笔记。

2.1.2 实验材料，设计和程序

主要的刺激事物是 5 列 18 个单词的字符串，是从 Yang 等人(2017)的实验 4 中获取的。按照拉丁方设计的方法确定呈现给参与者的单词的顺序。

参与者被告知他们要逐一学习 5 列单词，他们的任务是尽可能多地记住这些单词，最终会有一个总的测试。此外他们还被告知，每一列单词学习完之后，需要做 1 分钟的数学题，之后计算机会随机决定是否让他们回忆（即测试）刚刚学习的那列单词，不然的话就要再花一分钟的时间做数学演算题。实际上，临时测试是提前决定好了的。测试组每列测试做完就会做一个临时测试（interim test recall，即当列测试），而演算组则在前 4 列单词学习之后都做数学题，只有第五列单词学习之后才做一个当列测试。实验流程图见图 1。

在每列单词呈现时，单词是一个一个呈现在屏幕中央的，一个单词呈现 2 秒钟供参与者记忆，单词与单词之间间隔 500 毫秒。每列单词学完之后，两组参与者都要做一分钟的数学题（例如 45+62=？）。在临时测试时，参与者回忆当列单词的时间不做限制。5 列全部学习完毕后，两组都做一个最终的总体测试（final test recall），他们被要求对 5 列所有的单词尽可能多地进行回忆。所有的测试没有时间限制，也没有对错反馈。

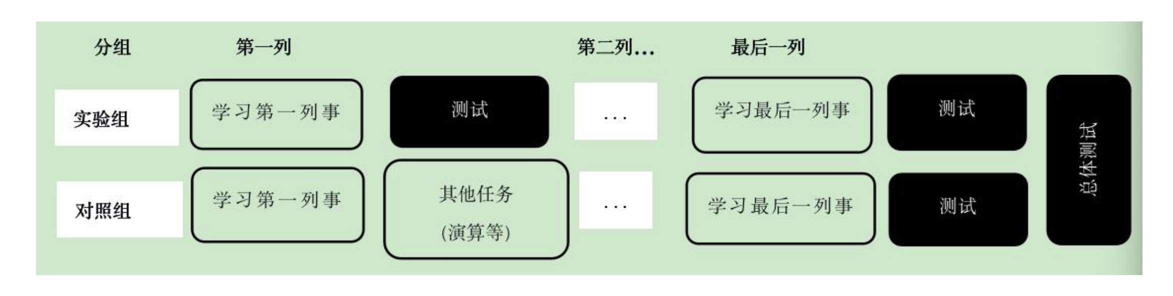


图 1 前测效应的实验流程

实验组（又称测试组）的每一列事物学习之后即进行当列测试，对照组（又称演算组或重学组）在前几列事物学习之后进行演算或者重新学习等其他任务。实验组和对照组最后一列学完后均进行当列测试，然后进行总体测试。

表 1 实验 1 实验组和对照组在控制变量上的组间差异

分组和组间差异	年龄（岁）	教育程度（年限）	健康(5 级)
测试组：均值（标准差）	66.73 (5.73)	15.13 (3.27)	3.53 (0.74)
演算组：均值（标准差）	66.13 (3.92)	15.92 (4.98)	3.33 (0.98)
差异：*t* 值（*p* 值）	0.33 (.74)	0.63 (.53)	0.30 (0.61)

注：均值（标准差）表示两组参与者人口学，其差异的显著性用 *t* 值和 *p* 值表示。以下各表相同。

2.1.3 实验结果

参与者人口统计信息的数据见表 1。两组之间的年龄、教育、健康水平方面没有明显的差别。

图 2A 展示了两组参与者临时测试的回忆成绩数目。对测试组组内重复测量的方差分析揭示出，在临时测试中对单词的正确回忆量随着列数的增加而显著减少，$F(1, 14) = 3.80$， $p = .07$，$\eta_p^2 = 0.21$，说明插入的临时测试没能完全阻止老年人随着学习列数的增加而对新的事物的回忆减少的趋势。最关键的是两组之间第 5 列当列测试的成绩差异。测试组($M = 5.87$, $SD = 1.21$)在第 5 列大约回忆出了两倍于演算组($M = 3.20$, $SD = 2.01$)的单词数目，平均相差 2.67 个单词，95% 置信空间为 [1.02, 4.32], $t(28) = 3.31$, $p = .003$, Cohen's $d = 1.21$,表明前置测试促进了老年人对新的事物的学习。

图 2B 展示了临时测试的前摄干扰效应（interim test PI，例如，当要求回忆某列单词时，参与者错误地回忆起了之前某列的单词）。对测试组第 2～5 列的组内方差分析显示，前摄干扰效应随着列数变多而加重，$F(1, 14) = 8.15$，$p = 0.01$，$\eta_p^2 = 0.37$，表明前置测试不能完全防止前摄干扰效应对老年人的影响。特别值得注意的是，演算组($M = 3.40$, $SD = 1.59$)出现了大约三倍于测试组($M = 1.20$, $SD = 1.21$)的前摄干扰效应，相差大约 2.2 个错误的单词，95% 置信空间为[1.14, 3.20]，$t(28) = 4.26$, $p < 0.001$，Cohen's $d = 1.56$，表明前置测试部分地阻止了前摄效应对老年人的干扰。

图 2C 展示了最终的回忆量。在最终的测试中，测试组的成绩(M = 19.73, SD = 7.24)显著地超过了演算组(M = 13.00, SD = 6.95), 平均相差 6.73 个单词, 95% 置信空间为 [1.43, 12.04], $t(28) = 2.60$, $p = 0.015$, Cohen's $d = 0.95$.

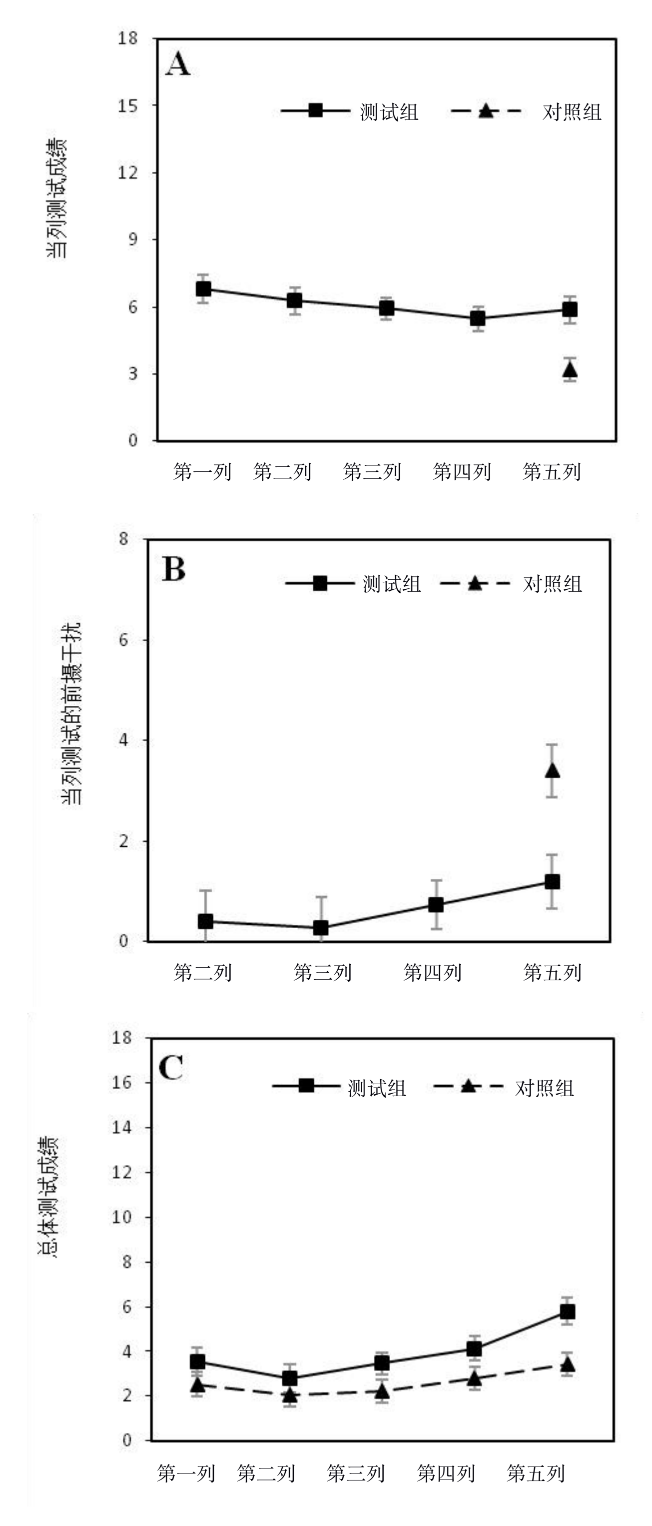


图 2 A: 临时测试在 5 列中的回忆成绩 B: 5 列临时测试中的前摄干扰 C: 最终测试中回忆起的每列的单词量(实验 1)

(误差条代表± 1 个标准差)

2.1.4 小结

尽管临时测试不能完全阻止老年人学习新的事物时随着数量的增多记忆而成绩逐下降的趋势，但是临时测试中的前置测试显著地增强了老年人学习和记忆新事物的效果，这表明学习新事物的前测效应同样适用于老年人。与前人有关年轻参与者的实验结果不同的是，临时测试不能完全消除前摄干扰效应对老年人的影响，尽管它们非常可观地阻止了前摄干扰效应对老年人记忆的损害。

2.2 实验 2

本实验在实验 1 研究设计的基础上进行了三个扩展。第一，实验 1 显示，相对于在列与列的学习之间插入演算，插入临时测试促进了老年人学习新事物的能力。实验 2 则在对照组的前 4 列学习之后插入重学而非演算，第 5 列与实验组同样是临时测试。第二，在实验 2 中，我们对参与者的基本认知能力进行了测试，以便确定在第 5 列临时测试时，实验组获得更好的回忆成绩是基于前 4 次的临时测试而非其他因素。第三，为了将前测效应实验的记忆任务更接近日常情境，学习材料分别尝试了面孔-姓氏配对，面孔-城市名配对，结果在预实验时都出现了地板效应，于是采用超市购物的事物，预实验表明不存在地板和天花板效应。

2.2.1 参与者

通过实验者熟悉的老人招募社区的老年参与者，主要在中国大陆武汉的两个居民小区内进行招募。样本量的大小取决于实验 1 的效应量（Cohen's $d = 1.21$），要观测到前测效应的显著水平($\alpha = 0.05$; power = 0.80) 至少需要每组 12 个样本。为了平衡所学习的事物出现的顺序，我们每个组设置了 15 个参与者。33 个 60 岁以上的老年人参与了本实验，有一个老年人在实验期间放弃了实验，最终获得了 32 位参与者全部的实验数据，其中有 24 位是女性。他们被随机分配入两组，测试组 16 名，重学组 16 名。有一个参与者报告曾经罹患过抑郁症（经过临床诊断），加入或排除她的数据不会改变整体结果的模式。所有的参与者母语为中文，他们逐一在安静的环境下进行测试，并且每个完成实验的老年人获得了 50 元人民币的补贴。

2.2.2 实验材料、设计和程序

从百度文库[4]中选取了 50 个常见超市商品的图片（例如苹果、牙刷、香波）作为主要的刺激项。这些项目被随机分配为 5 组，每组 10 个。每项图片下面用 2～3 个汉字指明商品的名称。采用拉丁方设计的方法平衡了实验材料对参与者呈现的顺序。

[4] 参见 https://wenku.baidu.com/view/9ef45d97915f804d2b16c1f0.html

在实验之初，所有的参与者报告了他们的年龄、性别和受教育年限。然后他们被要求报告是否罹患过精神疾病，如果是，还要报告精神疾病的种类。他们还被要求从 1（完全不好）到 5（非常好）分级评价一下他们自己的健康状况。

随后，实验者要求参与者们想像一下他们正准备去超市购物，随后会给他们提供五个购物清单，希望他们尽可能多地记住。其他的实验指导和程序与实验 1 一样。不同的是，本实验的对照组为重学组，在学习第 1 列到第 4 列的物品清单时，每学完一列，做 1 分钟的数学题，然后重学一遍（实验组为测试组，每学完一列，做 1 分钟的数学题后进行当列测试）。每个物品逐一呈现，每次呈现 2 秒。在完成最终测试之后，所有的参与者都要进行一个词汇认知能力的测试，并且通过数字广度任务测试其工作记忆能力（working memory capacities, WMC），以及通过认知加工任务测试其认知加工速度。

词汇任务

词汇任务来自于韦克斯勒成人智力量表(WALS; Wechsler, 1955)，在其汉化版本(龚耀先, 1992)里总共选取了 20 个二字词组成词汇表。实验者按照随机的顺序读出某个词汇，要求参与者在词汇表中找出该词并解释其意义。如果成功找到该词，得 1 分，能够正确地解释词意，再得 1 分。否则，得 0 分。满分 40 分。

数字广度任务

WMC 通过数字广度来进行测试(GrÉGoire & Van Der Linden, 1997)

认知加工速度任务

认知加工速度任务中，参与者被要求尽可能快而准确地判断两串数字是否完全一样(王大华，申继亮，彭华茂，唐丹，张凌，2005)。在每一次判断任务中，屏幕上会出现两串数字，每串数字由 9 个数字组成。如果 9 个数字完全相同，则要求参与者选择“一致”按键，如果有 1 个数字不同，则选择“不一致”按键。总共有 18 次任务判断。

表 2 实验 2 实验组和对照组在控制变量上的组间差异

分组和组间差异	年龄 (岁)	教育程度 (年限)	健康 (5 级)	词汇得分 (总分 40)	工作记忆	认知加工精度 (%)	认知加工速度(秒)
测试组：*均值* (*标准差*)	66.47 (4.93)	9.93 (3.28)	2.73 (0.46)	36.87 (3.31)	5.73 (0.98)	92.96 (7.99)	6.54 (1.11)
重学组：*均值* (*标准差*)	65.63 (4.24)	9.25 (2.29)	2.75 (0.77)	35.94 (3.15)	5.40 (0.95)	89.58 (10.32)	6.29 (1.81)
差异：*t* 值 (*p* 值)	0.51 (0.61)	0.68 (0.51)	-0.07 (0.94)	0.80 (.43)	0.94 (0.35)	1.02 (0.32)	0.46 (0.65)

注：均值（标准差）表示两组参与者人口学量值和基本认知能力，其差异的显著性用 *t* 值和 *p* 值表示。

2.2.3 实验结果

人口统计信息和基本认知能力的结果见表 2。两组之间的年龄、教育、健康水平、词汇能力、工作记忆、加工精度和速度等方面均没有显著的差异。

图 3A 展示了测试组的回忆成绩。对测试组组内重复测量的方差分析揭示出临时测试的正确回忆量随着列数的增加而显著减少，$F(1, 14) = 17.87$, $p = .001$, $\eta_p^2 = 0.56$. 测试组的参与者在第 5 列的当列测试中($M = 5.13$, $SD = 1.41$)回忆出了两倍于重学组($M = 2.38$, $SD = 1.63$)所回忆出来的物品量，平均相差 2.76 个物品, 95%置信空间为[1.64, 3.88], $t(29) = 5.03$, $p < .001$, Cohen’s $d = 1.80$.

图 3B 展示了临时测试的前摄干扰抑制效应。对测试组第 2～5 列的组内方差分析显示，前摄干扰效应随着列数变多而加重，$F(1, 14) = 20.70$, $p < 0.001$, $\eta_p^2 = 0.60$. 在第 5 列的临时测试中，重学组($M = 3.06$, $SD = 2.11$)也出现了大约三倍于测试组($M = 0.93$, $SD = 0.96$)的前摄干扰效应，相差大约 2.13 个物品, 95%置信空间为[0.91, 3.35], $t(29) = 3.57$, $p = 0.001$, Cohen’s $d = 1.30$.

图 3C 展示了最终的回忆量。在最终的测试中，测试组的成绩($M = 21.80$, $SD = 4.86$)显著地超过了重学组($M = 16.13$, $SD = 5.26$)，平均相差 5.68 个物品, 95%置信空间为[1.95, 9.45],

$t(29) = 3.11$, $p = 0.004$, Cohen's $d = 1.12$.

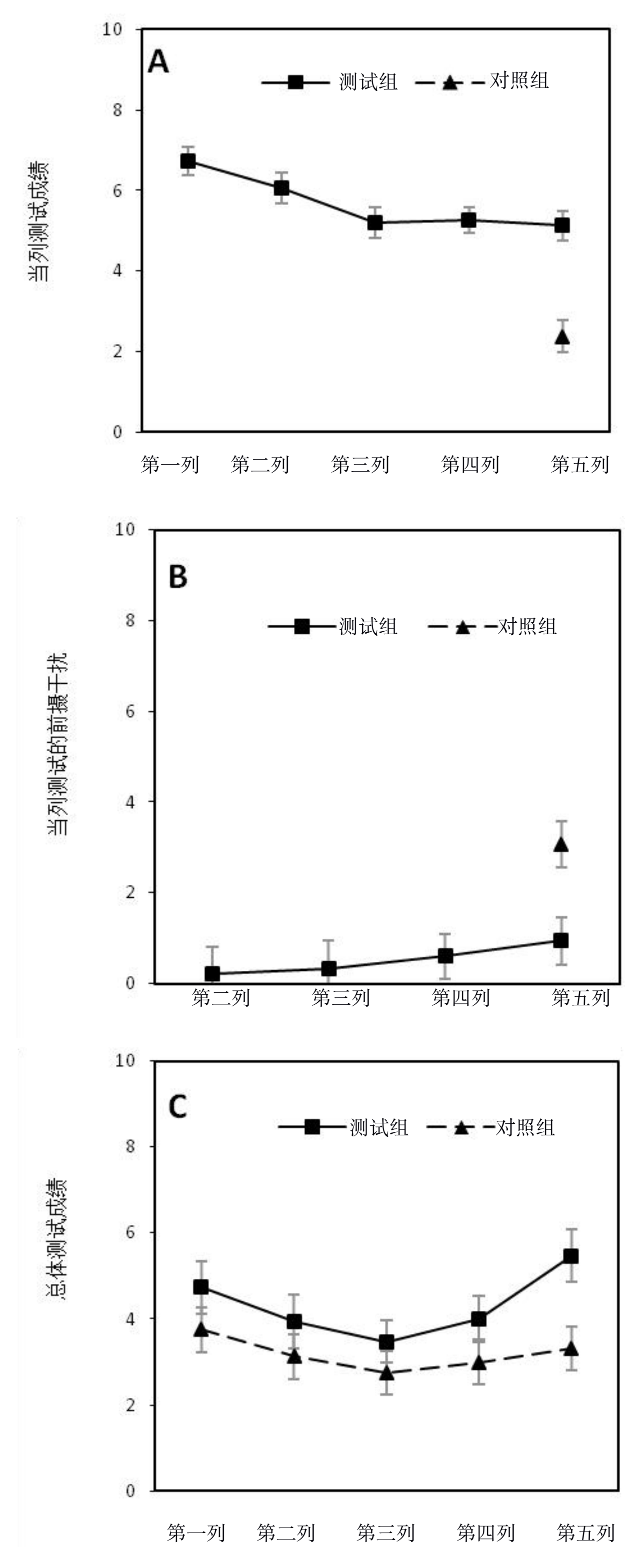


图 3 A:临时测试在 5 列中的回忆成绩 B:5 列临时测试中的前摄干扰 C:最终测试中回忆起的每列的单词量（实验 2）

（误差条代表± 1 个标准差）

2.2.4 小结

针对日常情境中的物品记忆而言，前置测验同样显著地增强了老年人学习和记忆新的事物的效果，比多学一遍的效果还要好。这又一次证明记忆的前测效应对老年人学习新事物有着重要的促进作用。当然，随着所要记忆的内容的增多，老年人的回忆成绩随着列数增多而显著下降，实验组插入的临时测试虽然部分阻止了前摄干扰的效应，但是无法阻止记忆效果逐列下降的总体趋势。

2.3 实验 3

实验 1 和 2 表明，前置测试能够提高老年人对新的简单事物的记忆，所以说老年人对简单事物的学习中存在前测效应。在实验 3 中，我们将检验前测效应是否可以推及到老年人对复杂事物的学习，从而验证老年人对复杂事物的学习是否会存在前测效应。

2.3.1 参与者

参与者主要来自于江汉油田老年大学，此外还在湖南湘潭市某社区进行了招募。共有 52 位老年人（年龄为 60 岁及以上）自愿报名参加。其中 3 人因未在 24 小时之内返回参加基础认知能力测试而放弃，余下的 49 人（38 名男性）参加测试。他们被随意分为 2 个小组，25 人组成实验组，24 人组成重学组（对照组）。所有的人均未参加实验 2。他们同样被测试了人口统计数据和基础认知能力。所有参与者的第一语言为汉语，分别在一隔音房间受试，测试完后每人获得 50 元人民币（或等值物品）的补贴。

2.3.2 实验资料，设计和程序

基本的刺激物为采集自 NetEase 公开课网站（http://open.163.com/movie/2015/2/s/8/MAH8PH40M MAKPPCAS8.html）关于医患关系的演讲视频。该演讲被分为 4 个段落，每段时长 4 分钟。为简单起见，每一段由 15 个短句构成。每一个句子含一个重点单词作为测试知识点。

样本的大小是根据 Szpunar 等人(2013)实验 2 中的效应量（Cohen's $d = 1.06$）决定的，样本量要求在每组的 24 个参与者中观察到显著的前测效应（$\alpha = 0.05$; power = 0.95）。在实验的初始阶段，受试者报告了人口统计信息。然后，指导他们学习 4 段演讲视频。告知他们第 4 段视频后要对所有 4 段视频进行最后测试，而且计算机会自行决定在学习完每一段后是否要给他们进行一段测试并做 30 秒的数学演算题。如果是，他们要就刚学习的视频段做填空测试；如果不是，他们要重新学习该视频段。事实上，实验组在学习每一段后都会进行临时试验，而控制组对 1～3 段进行了重新学习，只对第 4 段进行临时测试。在临时测试时，对刚才学习的段落的文字陈述出现在屏幕上，共包含 15 个句子，每一句留一空格（如：在医疗纠纷中，医生认为他们应首先____自己）。要求参试者在相应的空格中填写自己的答案。

对每一个问题的回答没有限定时间，如果不知道正确答案，允许不填空格。每次临时测试之后，参与者都会收到正确的反馈，并重新显示题目（在医疗纠纷中，医生认为在医疗纠纷中，他们应首先应保护自己）。正确答案用红色标示，并用下划线标注，参与者有 60 秒的时间学习正确反馈，逐段学习 1～3 段视频，并用 30 秒演算一些数学题。重学组则再次学习刚刚学过的视频，该段视频的文字陈述（含相同的 15 个句子）还要特别为他们显示 150 秒，15 个重要目标句子用红色标示，并标注下划线，提示参与者哪些单词会在最后测试时测试。重学组对第 4 段视频进行了与实验组相同的测试、反馈查看和数学演算。

完成第 4 段测试后，两组都进行了最后的测试，要求参与者回答 20 道填空题，1～4 段视频中各选 5 题。所有题目都显示于同一屏幕。我们只选择了 20 题，而非全部 60 题，因为在预研究时一些参与者告诉实验员他们希望缩短实验时间。参与者基本上都在 24 小时内返回并完成了认知能力的试验，这与实验 2 的参与者相同。

2.3.3 实验结果

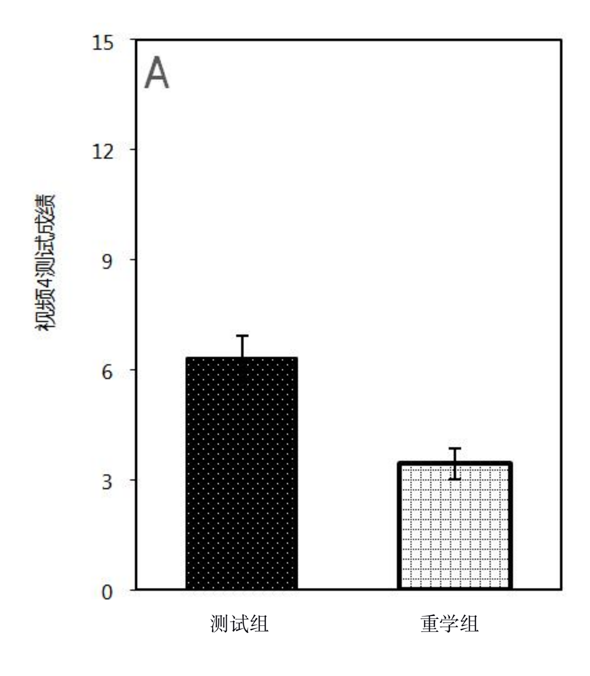
参与者人口统计信息和基本认知能力的结果见表 3。两组之间的年龄、教育、健康水平、词汇能力、工作记忆、加工精度和速度等方面没有明显的差别。两位对实验目的并不了解的实验助理对参与者的记忆测试进行了评分。一位实验助理对参与者所有的记忆测试进行评分，另一位实验助理只对第 4 段视频的临时测试进行评分。他们的评分（两位助理对所有参与者第 4 段视频的临时测试的评分）高度一致：他们评分的 95.9%是相同的。两位助理评分有差异的项目由实验者再次评分。

对于 1～3 段视频的测试，测试组的回忆率分别为 9.00（SD = 2.42）,9.08(SD = 3.09), 7.76(SD = 2.52)。两组参与者对第 4 段视频临时测试的成绩意义重大（参见图 4A）。在第 4 段视频的临时测试中，测试组（M = 6.32; SD = 3.11）的回忆率达到重学组（M = 3.50; SD = 2.17) 的两倍,平均相差 2.82 个项目，95%CI =〔1.28, 4.36〕,t (47) = 3.68，p = 0.001, Cohen's d = 1.05,由此可见前置测试可显著提高老年人对新的复杂事物的学习。如图 4B 所示，在最后总的测试中，测试组的成绩（M = 15.84；SD = 2.29）也优于重学组（M = 12.21; SD = 3.87)，其差别为 3.633 个项目，95%CI=〔1.81, 5.45〕,t（47）=4.02，p < 0.001, Cohen's d=1.14。

表 3 实验 3 实验组和对照组在控制变量上的组间差异

分组和组间差异	年龄 (岁)	教育程度 (年限)	健康 (5 级)	词汇得分 (总分 40)	工作记忆	认知加工精度 (%)	认知加工速度(秒)
测试组：均值 (标准差)	67.16 (6.44)	10.92 (3.11)	3.08 (0.64)	37.08 (3.21)	5.84 (1.07)	96.22 (5.25)	6.37 (2.13)
重学组：均值 (标准差)	66.96 (6.20)	11.04 (3.11)	3.21 (0.83)	36.21 (4.23)	5.73 (1.22)	96.30 (4.54)	5.73 (1.22)
差异：*t* 值 (*p* 值)	0.11 (0.91)	-0.14 (.89)	-0.61 (0.55)	0.81 (0.42)	0.34 (0.74)	-0.05 (0.96)	0.48 (0.64)

注：均值（标准差）表示两组参与者人口学量值和基本认知能力，其差异的显著性用 *t* 值和 *p* 值表示。



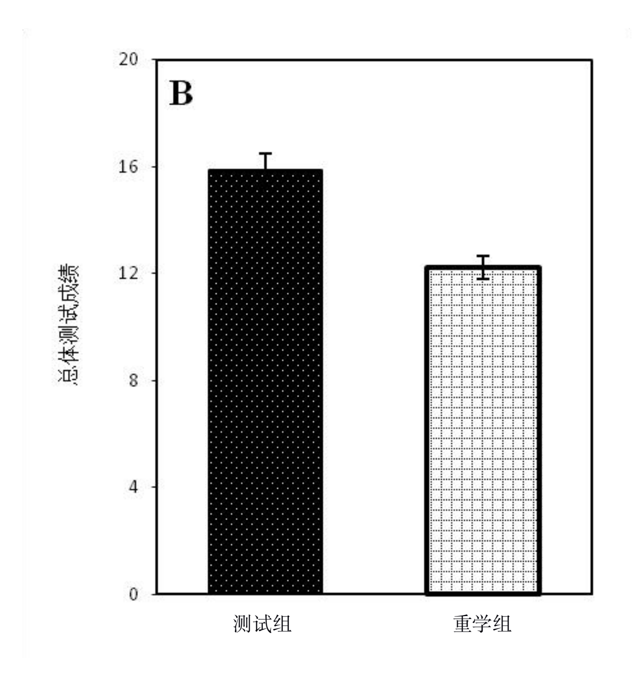


图 4 A:第 4 段视频临时测试成绩. B: 最终测试成绩（实验 3）

误差条代表 ±1 个标准差.

2.3.4 小结

该实验结果表明，关于复杂事物学习的前测效应也可以推及到老年人。由于复杂事物的学习不可能像简单事物的学习那样存在前摄干扰项目，所以实验 3 只比较了临时测试成绩和最终总测试成绩的差异。此外，令人耐人寻味的是，实验组第二段视频的临时测试成绩并不比第一段临时测试的成绩差，不过自第二段视频的学习之后，其后临时测试的正确率呈现出明显的不断下降的趋势，这与前两个实验是一致的。这说明老人的记忆负荷在实验设置中应当得到充分的考虑，且需要得到相应的理论解释。

3 结论与讨论

总而言之，无论是简单事物还是复杂事物，学习新事物的前测效应都可推及到老年人。产生这种结果的原因可能是前置测试引发了不同的心理作用机制，对于学习新的简单事物而言，可能主要是由于前置测试避免了前摄干扰（PI） 的建立，对于学习新的复杂事物而言，可能主要是由于前置测试在学习过程中缓解了学习参与的下降趋势，最终共同导致了老年人学习新事物过程中前测效应的出现。

具体而言，本研究实验 1 证明了在老年人学习简单材料的过程中进行临时测试，能有效降低前摄记忆项目的干扰作用——降低了大约三倍的前摄干扰效应。为了证明对于前摄干扰

效应的抑制不是由于临时测试的重复学习而带来的记忆效果, 实验 2 的对照组老年人在每一列记忆项目之后都进行了重复学习, 结果证明前置测试对老年人记忆新物品的效果仍然促进作用，测试组保持了相对于对照组抑制三倍前摄干扰效应的效果。

虽然近些年来前测效应被广泛印证, 但此前的研究大多局限于大学生 (Yang,et al., 2018；Chan, Meissner, & Davis，2018），前测效应是否可以推及 60 岁以上的老年人这一问题，在本研究展开时，没有任何针对这一特定年龄群体的报告见诸文献。本研究的主要目的是：验证是否可以利用前置测试作为弥补策略来提高老年人（60 岁以上）对新的信息的学习。如前文所述，对简单事物和复杂事物学习的前测效应可能通过两种机制来驱动：干扰降低和学习参与。实验 1 和实验 2 对老年人学习简单事物的前测效应的干扰降低机制进行了检验，发现干扰降低理论只能部分解释老年人学习新的简单事物的前测效应：本研究实验 1 和实验 2 的测试组记忆成绩显示，老年人即便是进行了前置测试，也无法阻止记忆成绩逐渐显著下降的趋势。而 Szpunar 等人（2008，实验 4）的研究显示，只要是插入了临时测试，前期的学习材料就不会对后期记忆的成绩有任何影响——每一列的正确回忆数量保持高度一致, 并且前摄干扰基本上不存在。这意味着对于实验 1 和实验 2 对照组中的老年参与者而言，其学习和记忆新事物出现成绩逐步下降的结果，其原因肯定不仅仅是前摄干扰，所以有理由推测，干扰降低理论无法完全解释老年人学习新事物时的前测促进作用。为了验证这一点，实验 3 将前测效应这一效果推及到老年人对复杂材料（演讲视频）的学习过程中，在这种被普遍认为前摄干扰有限、主要只存在测试组和重学组在记忆效果上的差别的情况下，测试组的回忆率几乎达到重学组的 2 倍, 这就进一步证实了还需要用其他的理论而非仅仅用干扰降低理论对老年人记忆前测效应进行有效解释。

本研究 3 个实验的前测效应不能完全用干扰降低理论来解释, 这很可能与参与者的年龄因素有关。根据 Aslan 和 Bäuml（2015）的发现，学习的前测效应无法推及到低龄儿童，原因在于对低龄儿童而言，无论有没有临时测试，都无法降低参与者学习和记忆新的事物的前摄干扰。虽然老年人与低龄儿童有所不同——前测效应也可以促进老年人对新事物的学习，但即便是插入了临时测试, 也无法阻止老人学习新事物时回忆成绩逐渐下降的趋势。很显然，干扰降低理论无法解释所有年龄阶段的学习者的前测效应，尤其是低龄儿童和老年人。

前人的研究普遍认为，对于复杂事物的学习而言，前摄干扰不会发生（Wissman, et al., 2011），所以关于学习复杂事物的前测效应的另一种解释是：前置测试可降低与任务无关的思绪游移（MW，Szpunar, et al.，2013）。有研究表明在老年人身上出现的 MW 比年轻人要少（Krawietz，et al.，2012），年龄的增长与 MW 的降低具有正相关（Seli, Maillet, Smilek,

Oakman, & Schacter, 2017）。本研究结果表明前置测试可以显著改善老年人对复杂事物的学习效果，所以就很难将复杂事物学习的前测效应仅仅归因于 MW 的下降。

此外，虽然有些研究发现老年人关于 MW 情节的报告的频率要低于年轻人，但是对于老年人而言，某些 MW 行为指数更高 (Zavagnin, Borella, & De Beni, 2014)。根据 Jordano 和 Touron（2017）的研究，老年人受到任务相关 MW 的影响要高于年轻人，他们发现老年人比年轻人更容易受到动机的影响，还发现其动机强度与无意 MW 呈负相关。由此可见，前置测试的促进作用很可能与老年人学习复杂事物的动机密切相关，即老人学习动机越强，则无意 MW 越低，同时任务相关 MW 会越高，本研究中的前置测试可能抑制了无意 MW，但更有可能是抑制了任务相关 MW。根据 Yang 等人（2019）的研究，在学习过程中前置测试能够改善参与者的学习参与状况，有利于参与程度的加强，本研究中老年人参与 3 个实验的记忆数据有力地支持了学习参与（LE）理论对前测效应的解释。

学习参与理论能够解释前测效应的诸多现象。第一，前置测试可以促使人们比控制组的人花更多的时间去对新信息进行编码（Yang, et al., 2017）。第二，Pastötter 等人（2011）发现：对于控制组的脑电波而言，在几个词列的学习过程中 α 波有明显的加强，而对于实验组则没有这种加强，这说明临时测试改变了所学事物的背景（目标搜索的范围），这就重置了对随后要学习的事物的记忆编码，使其编码的效率与前面的词列一样。第三，临时测试至少在短期内可以促进学习和改善记忆，甚至单元与单元之间事物的种类发生了改变时也有这种促进作用（Yang, et al., 2019），这就使得动机层面的解释更有说服力。

当然，学习参与理论对前测效应心理机制的解释还正在发展过程中。Weinstein 等人（2014）提出：临时测试可以使参与者产生对实验过程的预期，从而增强人们对新信息编码和提取的动机。Cho, Neely, Croco, 和 Vitrano（2017）提出：前置的临时测试中编码和提取的失误使人们产生不满意感，这促使参与者在后续临时测试中对编码和提取付出更多的努力。

综上所述，本研究结论认为，只有将干扰降低理论和学习参与理论结合起来，综合考虑参与者记忆能力的特点，才能更为有效地解释老年人学习和记忆过程中的前测效应，本文将之称为前测效应的“加减”理论。

首先，结合引言所述老年人记忆能力的特点和干扰降低理论，测试组老年参与者对新材料的记忆成绩（实验 1）较好，可能是因为前置测试的情节记忆减少了回忆新事物的背景项，所以前摄干扰的降低部分促进了老年人学习新事物的前测效应，但是由于老年人的情节记忆普遍有所下降，所以即便是插入临时测试，也无法完全阻止老年人学习新事物时回忆成绩逐

列下降的总体趋势。也就是说，老年人情节记忆能力的普遍“减”退和对学习新事物的干扰项的“减”消共同影响着实验 1 老人最终的记忆成绩。

其次，虽然老年人的未来记忆有所下降，但是前置测试可能简化了他们对目标事物进行编码和提取的背景，未来记忆的负荷降低，增强了他们学习参与的信心和动机，最终改善了他们日常生活中的未来记忆(试验 2)。但是由于老年人未来记忆的能力普遍有所下降，所以即便是插入临时测试，也无法完全阻止老年人记忆新事物时回忆成绩逐列下降的总体趋势。也就是说，老年人未来记忆能力的普遍“减”退、记忆背景项的“减”少和学习参与动机的增“加”共同影响着实验 2 老人最终的记忆成绩。

最后，只有将干扰降低理论和学习参与理论结合起来，才能充分解释老年人学习复杂材料过程中的前测效应（实验 3）。按照干扰降低理论，学习复杂材料的前测效应是因为思绪游移的降低，但是如前所述，老年人与任务无关的思绪游移较少而与任务相关的思绪游移较多，这使得干扰降低理论的解释遇到了困境。不过，根据 Jordano 和 Touron（2017）的研究发现，与任务相关的思绪游移的触发因素为与年龄相关的刻板印象，这些有关老化的消极刻板印象使老年人将注意力集中于有关自我评价威胁的学习环境。在实验 3 中，一方面临时测试的“随机性”降低了老年人面临自我评价的威胁，从而降低了与任务相关的思绪游移，另一方面前置测试又激发了他们的学习参与性，因而测试成绩要比仅仅伴随着与任务相关的思绪游移的状态时更好。也就是说，前置测试一方面促进了老年人在记忆过程中的学习参与，另一方面抑制了对新事物学习具有干扰作用的思维游移，从而共同提高了对新的复杂事物的记忆成绩。

根据本文提出的“加减”理论可以推测，文化因素、老年刻板印象的启动等因素都会影响到临时测试对老年人学习新事物的效果。Levy 和 Langer（1994）的研究发现，对老年人消极的刻板印象影响着老年人的记忆成绩。由于中国文化中有关老化的消极刻板印象较少，所以记忆成绩比美国的老年人要好。本研究限于篇幅，虽然实验参与者包括了不同文化下的老年人，但是没有深入探讨文化和自我评价的因素（老年刻板印象）对前测效应的影响，这将会在以后的研究中进一步展开。

本研究对于老年人学习和记忆的前测效应的研究推进，有望用来改善老年人因年龄增长而产生的记忆下降问题。有研究者(Kielar, Shah-Basak, Deschamps, Jokel, & Meltzer, 2019)发现原发性进行性失语症患者失去了说话或理解语言的能力，是因为大脑的某些区域出现了功能缺陷，而这些区域在核磁共振（MRI）影像学上尚未显示出结构损伤，所以推断可能是大脑功能性的缺陷而非结构性的障碍导致了老年痴呆症状。也就是说，老年痴呆患者根源在

于日常生活中功能发挥的失常而非大脑结构性病变，本研究给该问题带来了新的解决途径：将阶段性回溯的学习模式作为老年人的记忆训练策略，弥补他们的认知功能劣势，从而有望延缓乃至预防老年痴呆症的发生和发展。

5 致谢

感谢在本文研究中积极参与的所有老人和实验助理，包括但不限于王艳、杨帆、成方琪、刘宇晋、刘重阳等人。感谢评审专家对本文写作提出的每一个意见和建议，他们的悉心指导使得本文的质量得到极大的提升。

6 参考文献

Aslan, A., & Bäuml, K. H. T. (2015). Testing enhances subsequent learning in older but not in younger elementary school children. *Developmental Science, 19*(6), 992–998.

Chan, J. , Meissner, C. , & Davis, S.（2018）.Retrieval potentiates new learning： A theoretical and meta-analytic review.*Psychological Bulletin,144*(11), 1111–1146.

Cho, K. W., Neely, J. H., Crocco, S., & Vitrano, D. (2017). Testing enhances both encoding and retrieval for both tested and untested items. *Quarterly Journal of Experimental Psychology, 70*(7), 1211–1235.

Craik, F. I. M., & Rose, N. S. (2012). Memory encoding and aging: A neurocognitive perspective. *Neuroscience and Biobehavioral Reviews, 36*(7), 1729–1739.

Farrimond, S. J., Knight, R. G., & Titov, N. (2003). Prospective remembering and aging. *Archives of Clinical Neuropsychology, 18*(7), 769–769.

Gluck, J., & Bluck, S. (2007). Looking back across the life span: A life story account of the reminiscence bump. *Memory and Cognition, 35*(8), 1928–1939.

Gong,Y.(1992).*Revised Wechsler adult intelligence scale manual in China.* Hunan Map Press.[龚耀先（1992). *中国修订韦氏成人智力量表手册*. 湖南地图出版社]

GrÉGoire, J., & Van Der Linden, M. (1997). Effect of age on forward and backward digit spans. *Aging, Neuropsychology, and Cognition, 4*(2), 140–149.

Hafkemeijer, A., Van, D. G. J., & Rombouts, S. A. (2012). Imaging the default mode network in aging and dementia. *Biochimica et Biophysica Acta,1822*(3), 431–441.

Ikier, S., & Hasher, L. (2006). Age differences in implicit interference. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 61*(5), 278–284.

Ikier, S., Yang, L., & Hasher, L. (2008). Implicit proactive interference, age, and automatic versus controlled retrieval strategies. *Psychological Science, 19*(5), 456–461.

Jackson, J. D., & Balota, D. A. (2012). Mind-wandering in younger and older adults: Converging evidence from the sustained attention to response task and reading for comprehension. *Psychology and Aging, 27*(1), 106–119.

Jing, H. G., Szpunar, K. K., & Schacter, D. L. (2016). Interpolated testing influences focused attention and improves integration of information during a video-recorded lecture. *Journal of Experimental Psychology: Applied, 22*(3), 305–318.

Jordano, M., L., & Touron, D. R. (2017). Stereotype threat as a trigger of mind-wandering in older adults. *Psychology and Aging, 32*(3), 307–313.

Kielar, A. , Shah-Basak, P. P. , Deschamps, T. , Jokel, R. , & Meltzer, J. . (2019). Slowing is slowing: Delayed neural responses to words are linked to abnormally slow resting state activity in primary progressive aphasia. *Neuropsychologia, 126*(6),331–347.

Krawietz, S. A., Tamplin, A. K., & Radvansky, G. A. (2012). Aging and mind wandering during text comprehension. *Psychology and Aging, 27*(4), 951–958.

Levy, B. , & Langer, E. . (1994). Aging free from negative stereotypes: Successful memory in china among the american deaf. *Journal of Personality and Social Psychology, 66*(6), 989–997.

Nunes, L. D., & Weinstein, Y. (2012). Testing improves true recall and protects against the build-up of proactive interference without increasing false recall. *Memory, 20*(2), 138–154.

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging, 23*(1), 104–118.

Pastötter, B., & Bäuml, K. H. (2014). Retrieval practice enhances new learning: the forward effect of testing. *Frontiers in Psychology, 5*(4),1–5.

Pastötter, B., & Bäuml, K.-H.(2019). Testing enhances subsequent learning in older adults. *Psychology and Aging,34*(2), 242–250.

Pastötter, B., Schicker, S., Niedernhuber, J., & Bäuml, K. H. (2011). Retrieval during learning facilitates subsequent memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(2), 287–297.

Pastötter, B., Weber, J., & Bäuml, K. H. (2013). Using testing to improve learning after severe traumatic brain injury. *Neuropsychology, 27*(2), 280–285.

Piolino, P., Desgranges, B., Benali, K., & Eustache, F. (2002). Episodic and semantic remote autobiographical memory in ageing. *Memory, 10*(4), 239–257.

Schacter, D. L., & Szpunar, K. K. (2015). Enhancing attention and memory during video-recorded lectures. *Scholarship of Teaching and Learning in Psychology, 1*(1), 60–71.

Seli, P., Maillet, D., Smilek, D., Oakman, J. M., & Schacter, D. L. (2017). Cognitive aging and the distinction between intentional and unintentional mind wandering. *Psychology and Aging, 32*(4), 315–324.

Szpunar, K. K., Jing, H. G., & Schacter, D. L. (2014). Overcoming overconfidence in learning from video-recorded lectures: Implications of interpolated testing for online education. *Journal of Applied Research in Memory and Cognition, 3*(3), 161–164.

Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *Proceedings of the National Academy of Sciences of the United States of America, 110*(16), 6313–6317.

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1392–1399.

Wang, D., Shen, J., Peng, H., Tang, D., & Zhang, L. The model of educational effect on old adult's cognition. *Acta Psychologica Sinica.37*(4), 511–516.[王大华，申继亮，彭华茂，唐丹, & 张凌. (2005). 教育水平对老年人认知能力的影响模式. *心理学报, 37*(4), 511–516]

Wechsler, D. (1955). *Manual for the Wechsler adult intelligence scale.* Oxford, England: Psychological Corp.

Weinstein, Y., Gilmore, A. W., Szpunar, K. K., & McDermott, K. B. (2014). The role of test expectancy in the build-up of proactive interference in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(4), 1039–1048.

Weinstein, Y., McDermott, K. B., & Szpunar, K. K. (2011). Testing protects against proactive interference in face-name learning. *Psychonomic Bulletin and Review, 18*(3), 518–523.

Whitebourne, S. K., & Whitebourne, S. B. (2014). *Adult development and aging: Biopsychosocial perspectives (5th ed.)*. USA: John Willey & Sons.

Wiggs, C. L., Weisberg, J., & Martin, A. (2006). Repetition priming across the adult lifespan – The long and short of it. *Aging, Neuropsychology And Cognition, 13*(3-4), 308–325.

Wissman, K. T., Rawson, K. A., & Pyc, M. A. (2011). The interim test effect: Testing prior material can facilitate the learning of new material. *Psychonomic Bulletin and Review, 18*(6), 1140–1147.

Yang, C., Chew, S., Sun, B., & Shanks, D. R. (2019). The forward effects of testing transfer to different domains of learning. *Journal of Educational Psychology,111*(5),809–826.

Yang, C., Potts, R., & Shanks, D. R. (2017). The forward testing effect on self-regulated study time allocation and metamemory monitoring. *Journal of Experimental Psychology: Applied, 23*(3), 263–277.

Yang, C. , Potts, R. , & Shanks, D. R. . (2018). Enhancing learning and retrieval of new information: A review of the forward testing effect. *Npj Scientific of Learning, 3*(8),1–9.

Yang, C. , & Shanks, D. R. . (2018). The forward testing effect: interim testing enhances inductive learning. *Journal of Experimental Psychology: Learning, Memory & Cognition*,44(3), 485–492.

Zavagnin, M., Borella, E., & De Beni, R. (2014). When the mind wanders: Age-related differences between young and older adults. *Acta Psychologica, 145,* 54–64.